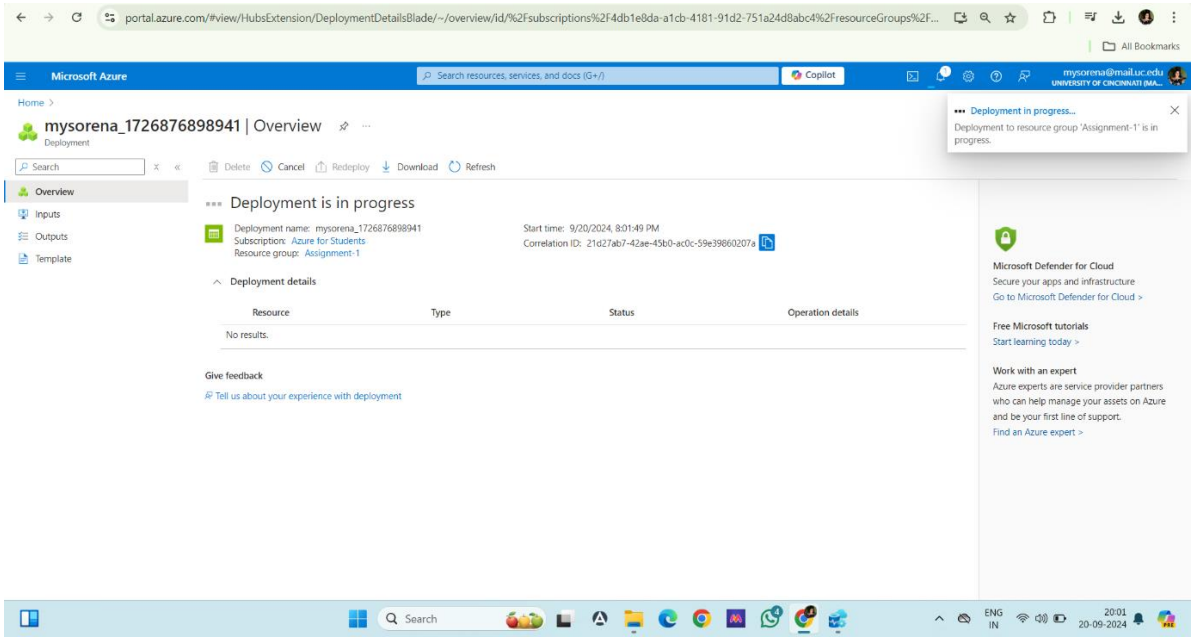
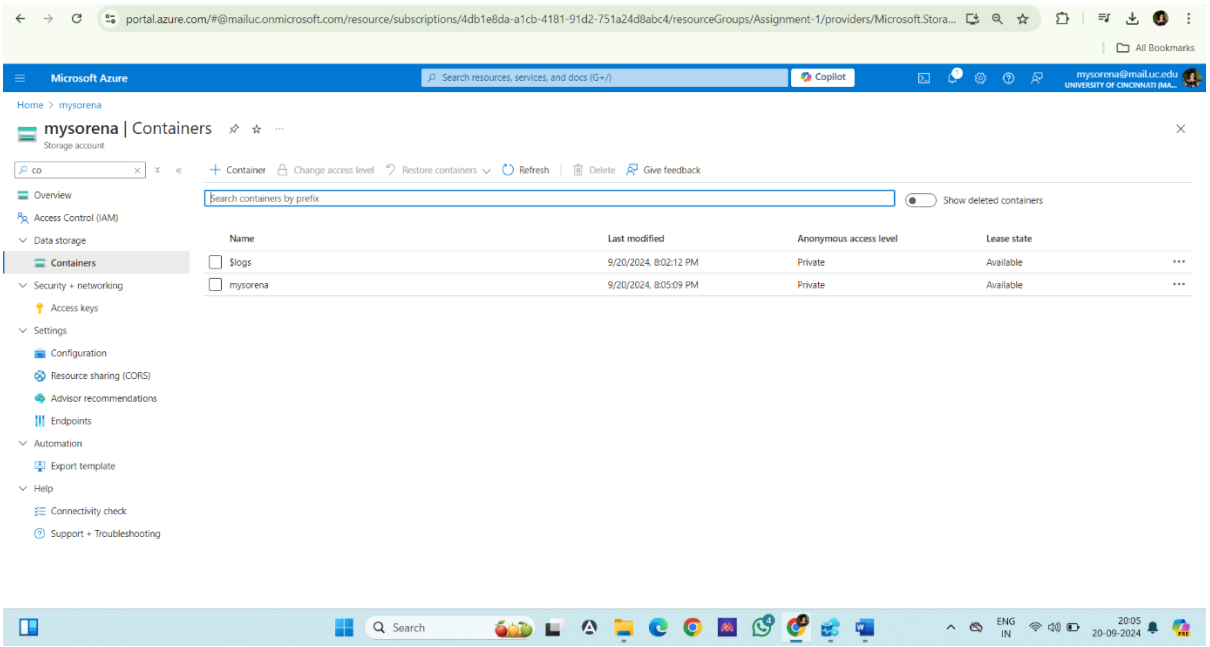
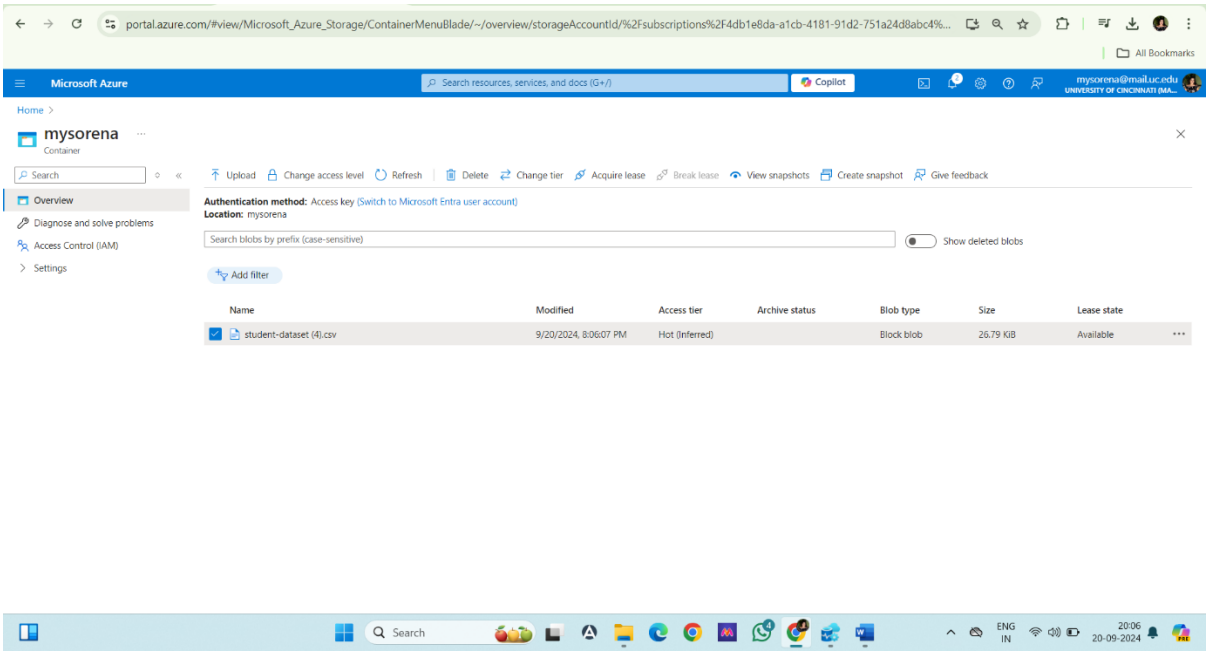
Cloud Computing Assignment

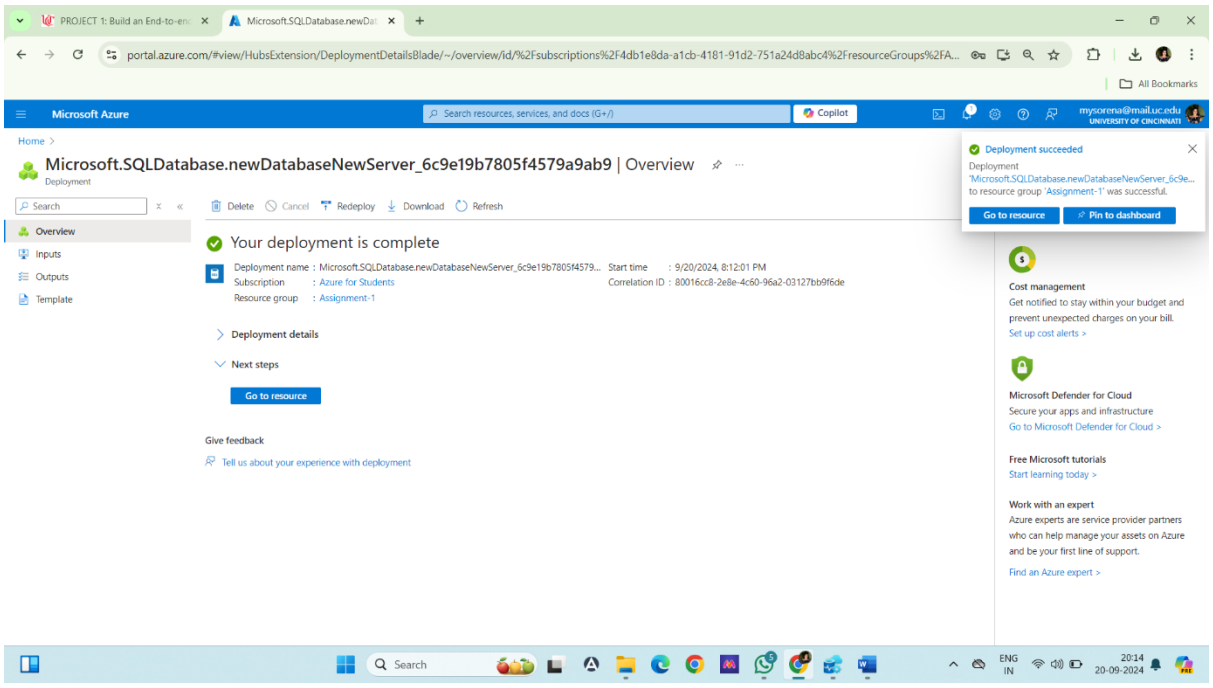
Set up an Azure Storage account and upload the student.csv

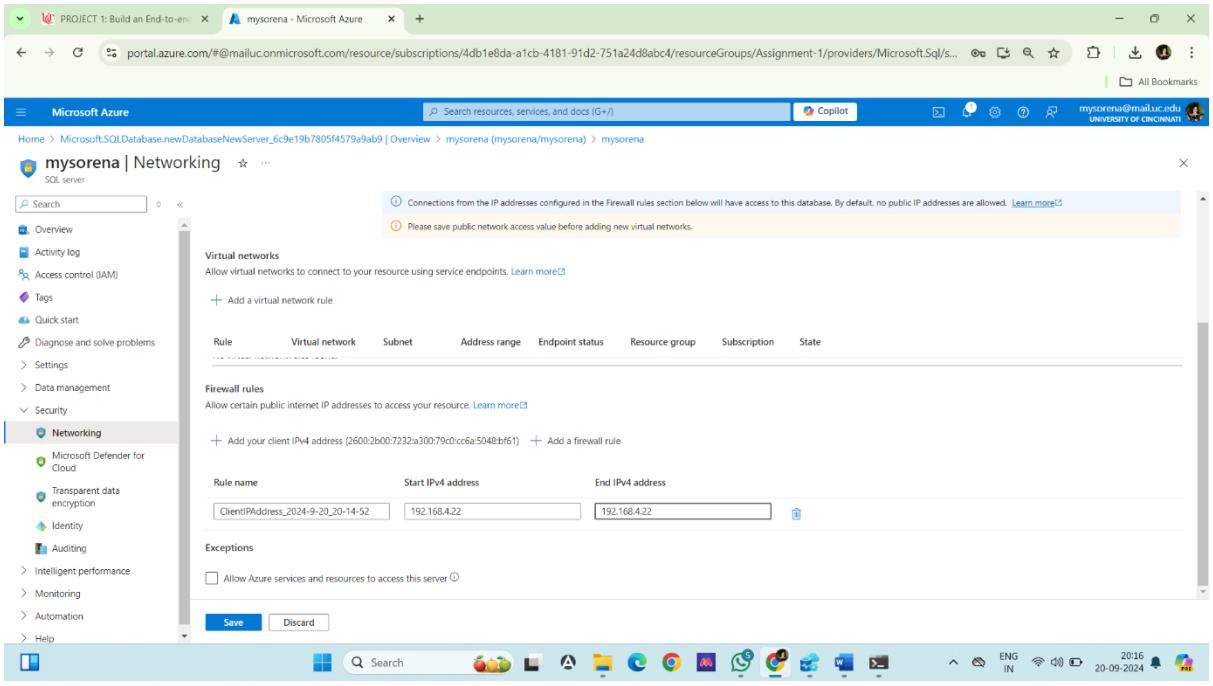


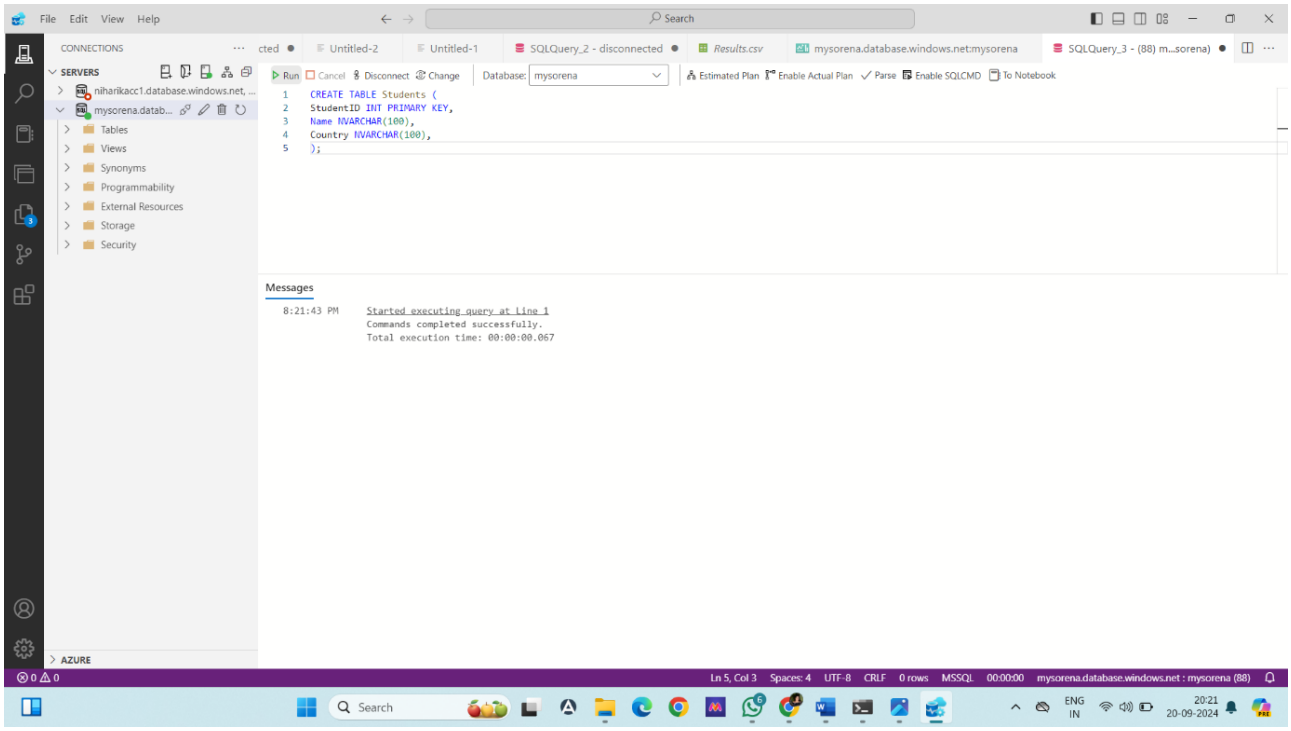


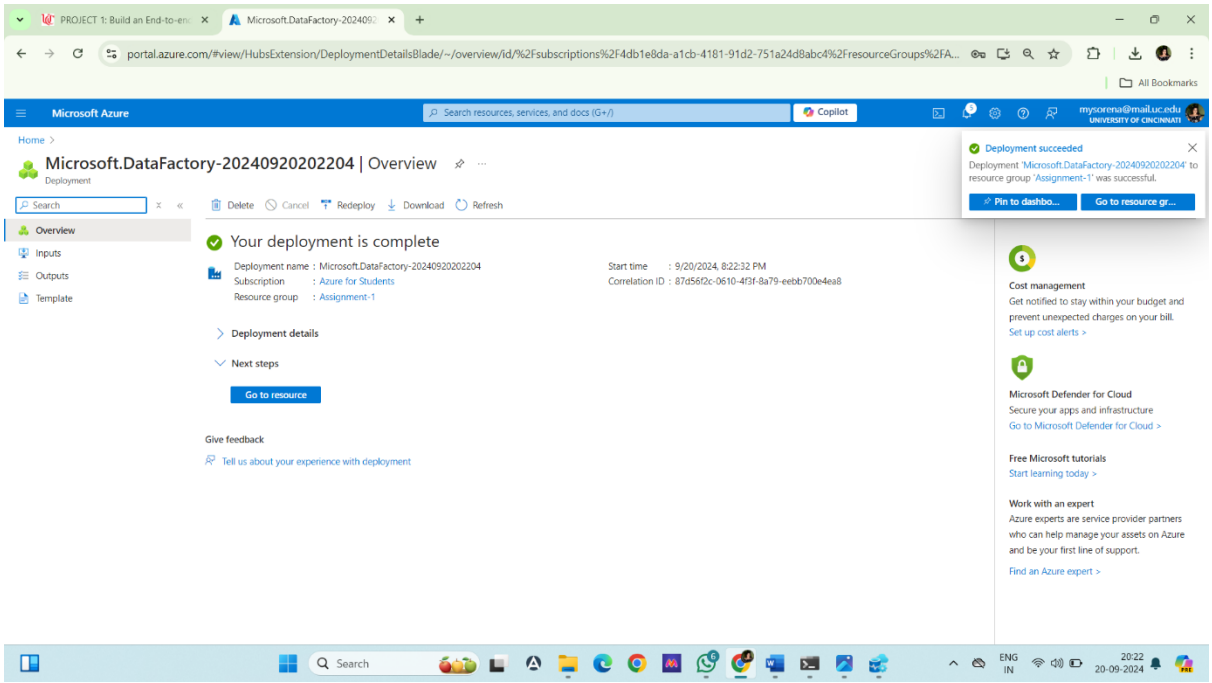

portal.azure.com/#view/Microsoft_Azure_Storage/ContainerMenuBlade/~/overview/storageAccountId/%2Fsubscriptions%2F4db1e8da-a1cb-4181-91d2-751a24d8abc4%...
All Bookmarks
Microsoft Azure
Search resources, services, and docs (G+/)
Copilot
mysorena@mail.uc.edu
UNIVERSITY OF CINCINNATI (MA...
Home
mysorena
Container
Search
Overview
Diagnose and solve problems
Access Control (IAM)
Settings
Upload
Change access level
Refresh
Delete
Change tier
Acquire lease
Break lease
View snapshots
Create snapshot
Give feedback
Authentication method: Access key (Switch to Microsoft Entra user account)
Location: mysorena
Search blobs by prefix (case-sensitive)
Show deleted blobs
Add filter
Name
Modified
Access tier
Archive status
Blob type
Size
Lease state
student-dataset (4).csv
9/20/2024, 8:06:07 PM
Hot (Inferred)
Block blob
26.79 KiB
Available
Search
ENG
IN
20:06
20-09-2024

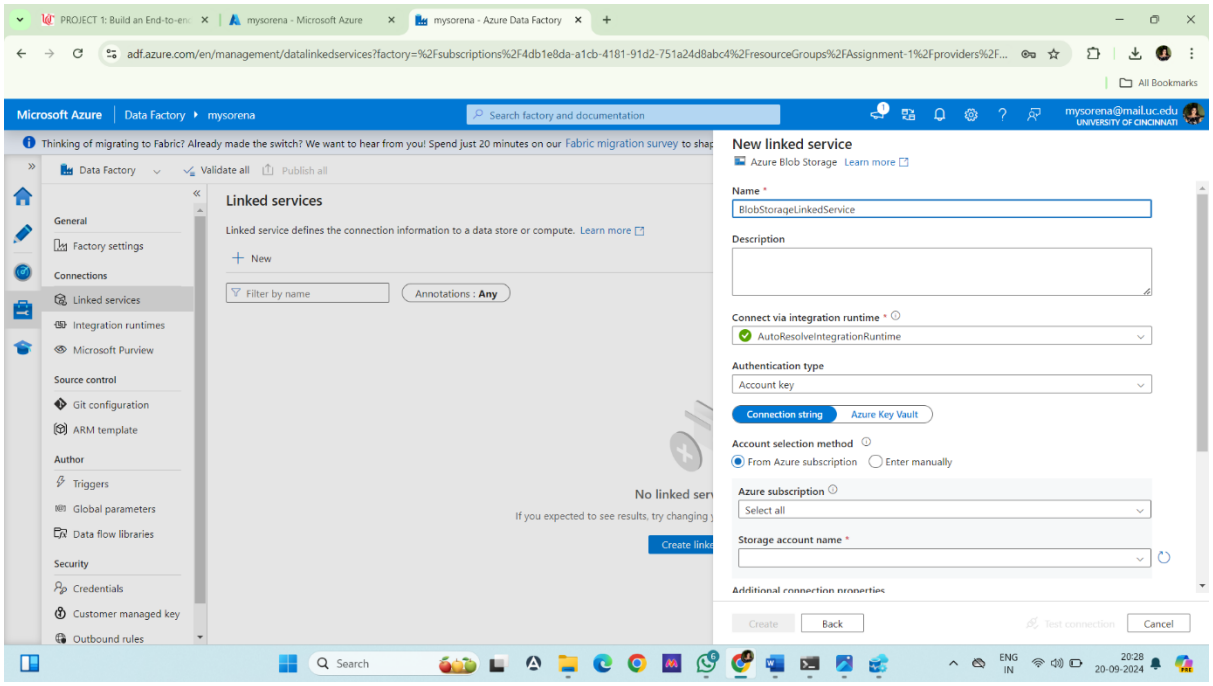
Set up an Azure SQL Database and create the Students table

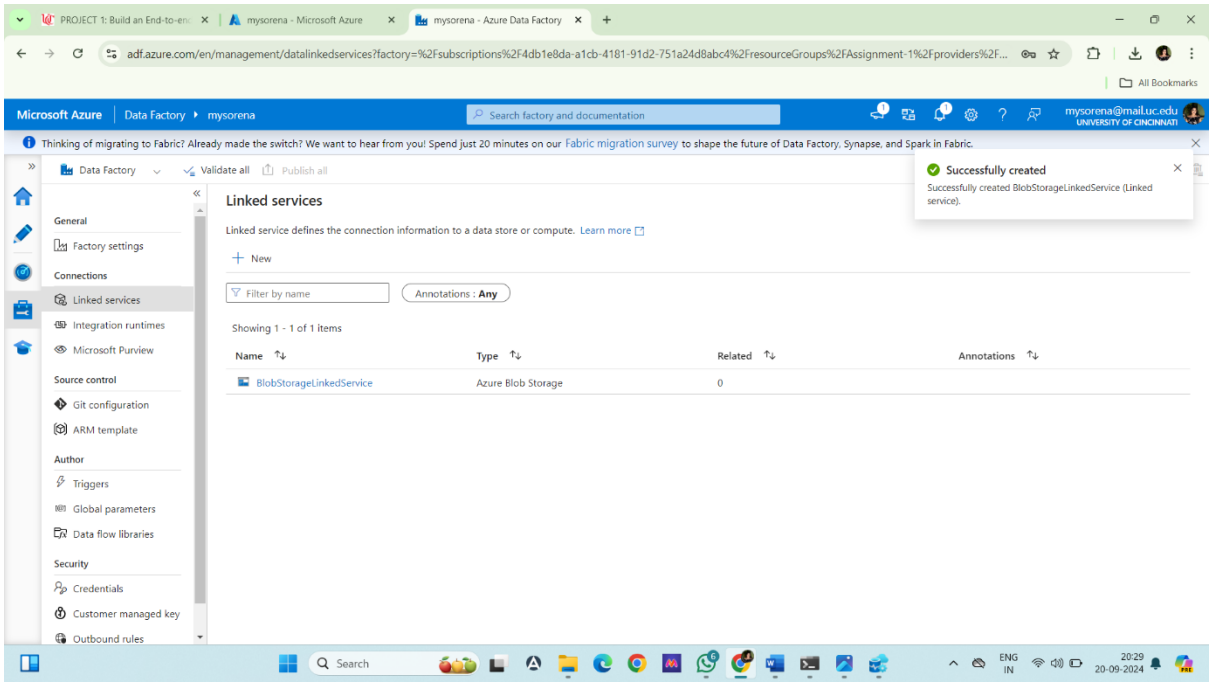


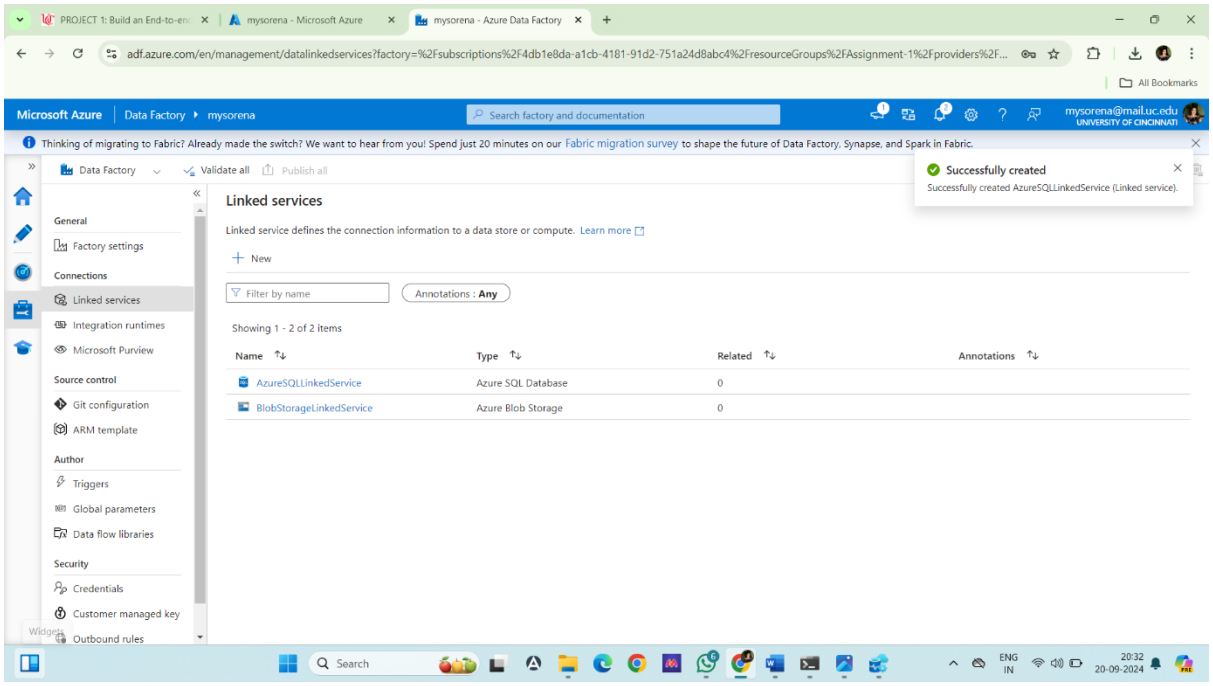


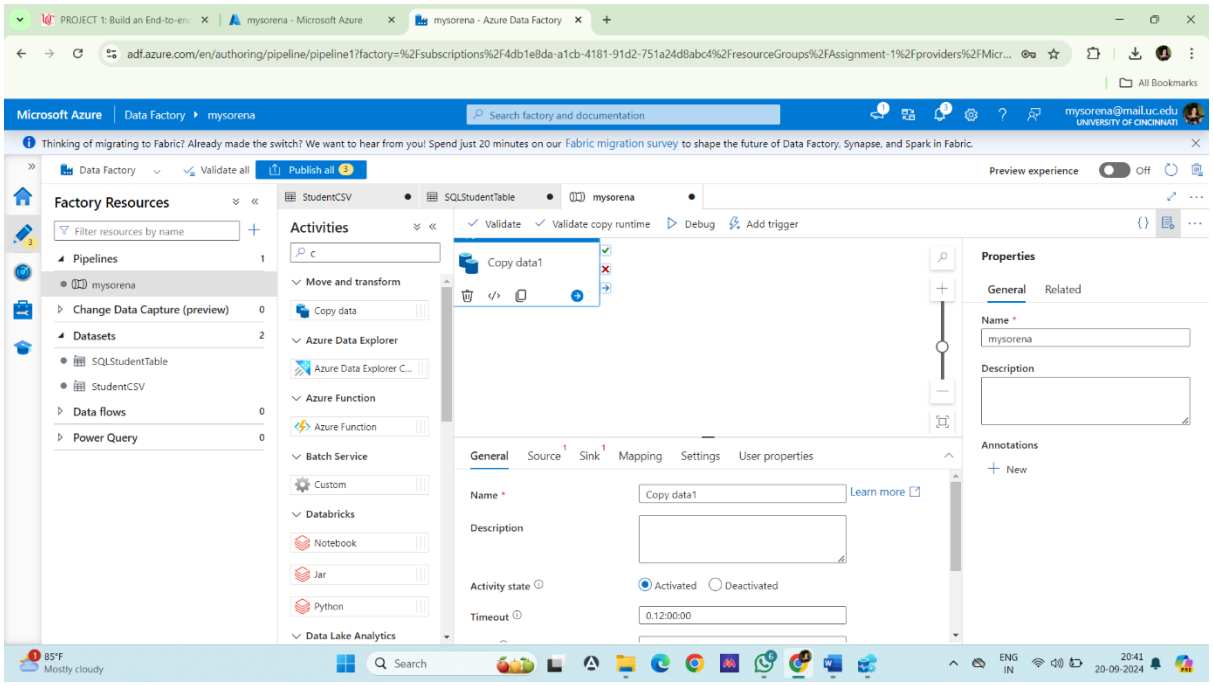

File Edit View Help
Search
CONNECTIONS
SERVERS
niharikacc1.database.windows.net, ...
mysorena.datab...
Tables
Views
Synonyms
Programmability
External Resources
Storage
Security
AZURE
Untitled-2
Untitled-1
SQLQuery_2 - disconnected
Results.csv
mysorena.database.windows.net:mysorena
SQLQuery_3 - (88) m...sorena)
Run Cancel Disconnect Change
Database: mysorena
Estimated Plan Enable Actual Plan Parse Enable SQLCMD To Notebook
1 CREATE TABLE Students (
2 StudentID INT PRIMARY KEY,
3 Name NVARCHAR(100),
4 Country NVARCHAR(100),
5);
Messages
8:21:43 PM Started executing query at Line 1
Commands completed successfully.
Total execution time: 00:00:00.067
Ln 5, Col 3 Spaces: 4 UTF-8 CRLF 0 rows MSSQL 00:00:00 mysorena.database.windows.net : mysorena (88)
Search
ENG IN
20:21
20-09-2024

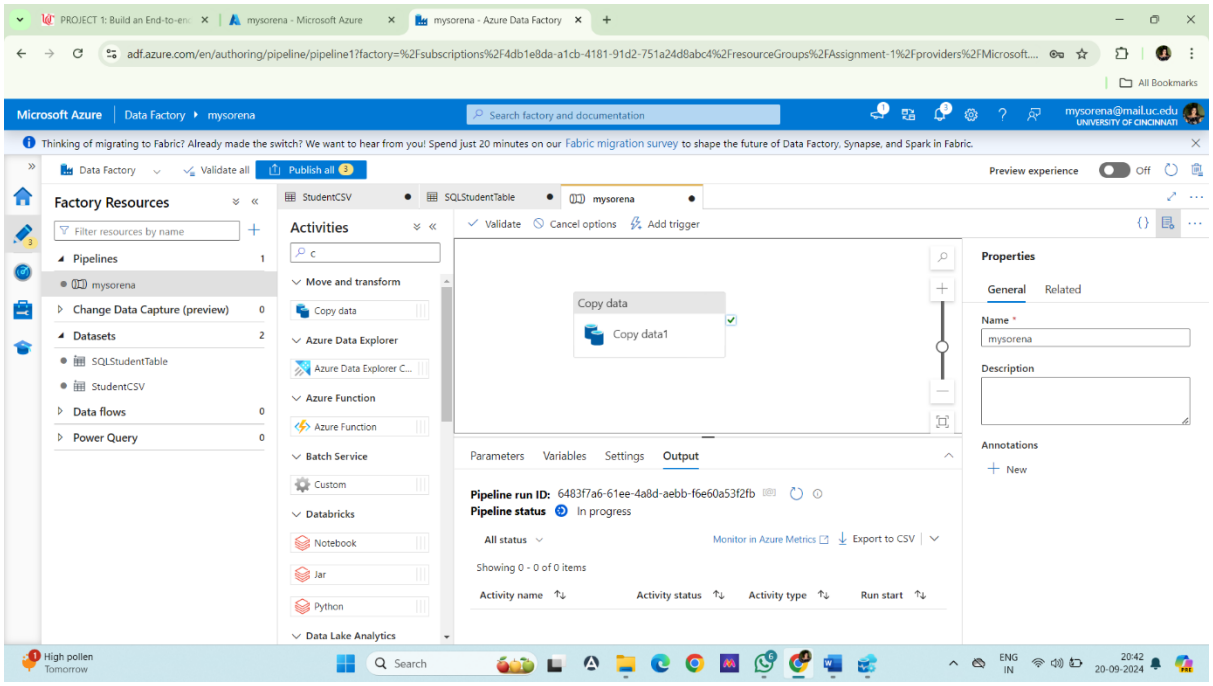

PROJECT 1: Build an End-to-end
Microsoft.DataFactory-20240920...
portal.azure.com/#view/HubsExtension/DeploymentDetailsBlade/~/overview/id/%2Fsubscriptions%2F4db1e8da-a1cb-4181-91d2-751a24d8abc4%2FresourceGroups%2FA...
All Bookmarks
Microsoft Azure
Search resources, services, and docs (G+/)
Copilot
mysorena@mail.uc.edu
UNIVERSITY OF CINCINNATI
Home
Microsoft.DataFactory-20240920202204 | Overview
Deployment
Search
Overview
Inputs
Outputs
Template
Delete Cancel Redeploy Download Refresh
Your deployment is complete
Deployment name : Microsoft.DataFactory-20240920202204
Subscription : Azure for Students
Resource group : Assignment-1
Start time : 9/20/2024, 8:22:32 PM
Correlation ID : 87d56f2c-0610-4f3f-8a79-eebb700e4ea8
Deployment details
Next steps
Go to resource
Give feedback
Tell us about your experience with deployment
Deployment succeeded
Deployment 'Microsoft.DataFactory-20240920202204' to resource group 'Assignment-1' was successful.
Pin to dashbo...
Go to resource gr...
Cost management
Get notified to stay within your budget and prevent unexpected charges on your bill.
Set up cost alerts >
Microsoft Defender for Cloud
Secure your apps and infrastructure
Go to Microsoft Defender for Cloud >
Free Microsoft tutorials
Start learning today >
Work with an expert
Azure experts are service provider partners who can help manage your assets on Azure and be your first line of support.
Find an Azure expert >
Search
ENG IN
20:22
20-09-2024

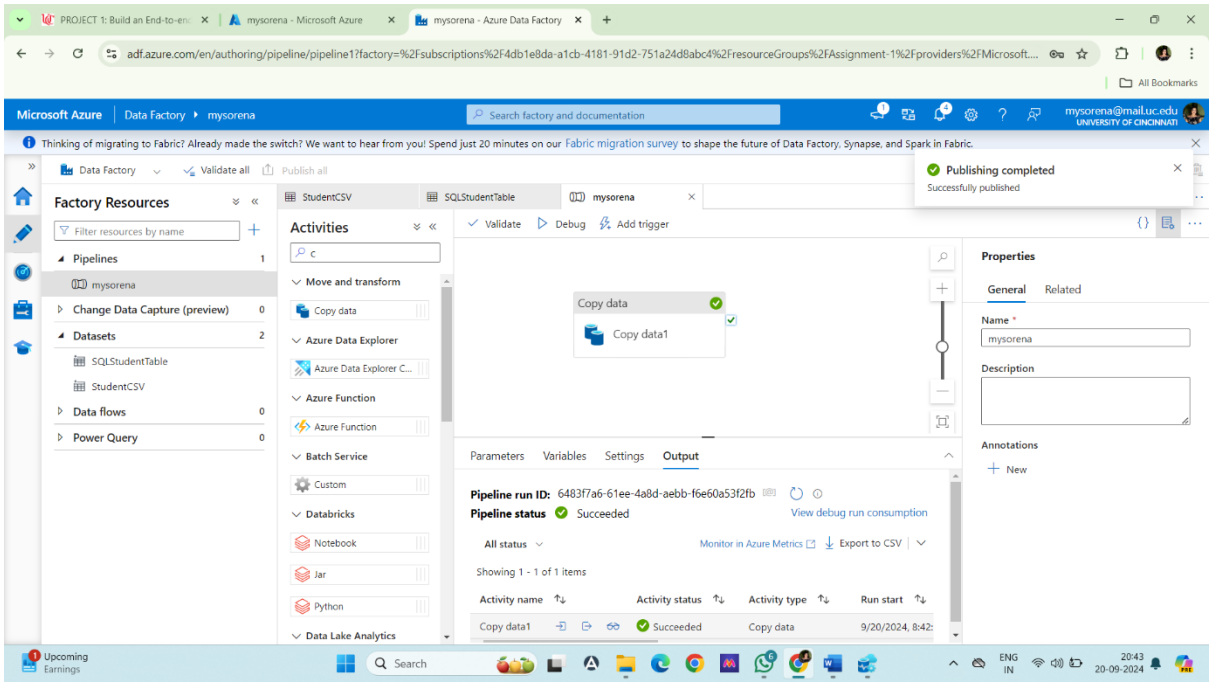

Microsoft Azure
Data Factory
mysorena
Search factory and documentation
mysorena@mail.uc.edu
UNIVERSITY OF CINCINNATI
Thinking of migrating to Fabric? Already made the switch? We want to hear from you! Spend just 20 minutes on our Fabric migration survey to shap
Data Factory
Validate all
Publish all
General
Factory settings
Connections
Linked services
Integration runtimes
Microsoft Purview
Source control
Git configuration
ARM template
Author
Triggers
Global parameters
Data flow libraries
Security
Credentials
Customer managed key
Outbound rules
Linked services
Linked service defines the connection information to a data store or compute. Learn more
New
Filter by name
Annotations : Any
No linked ser
If you expected to see results, try changing
Create link
New linked service
Azure Blob Storage Learn more
Name *
BlobStorageLinkedService
Description
Connect via integration runtime *
AutoResolveIntegrationRuntime
Authentication type
Account key
Connection string
Azure Key Vault
Account selection method
From Azure subscription
Enter manually
Azure subscription
Select all
Storage account name *
Additional connection properties
Create
Back
Test connection
Cancel
Search
ENG
IN
20:28
20-09-2024

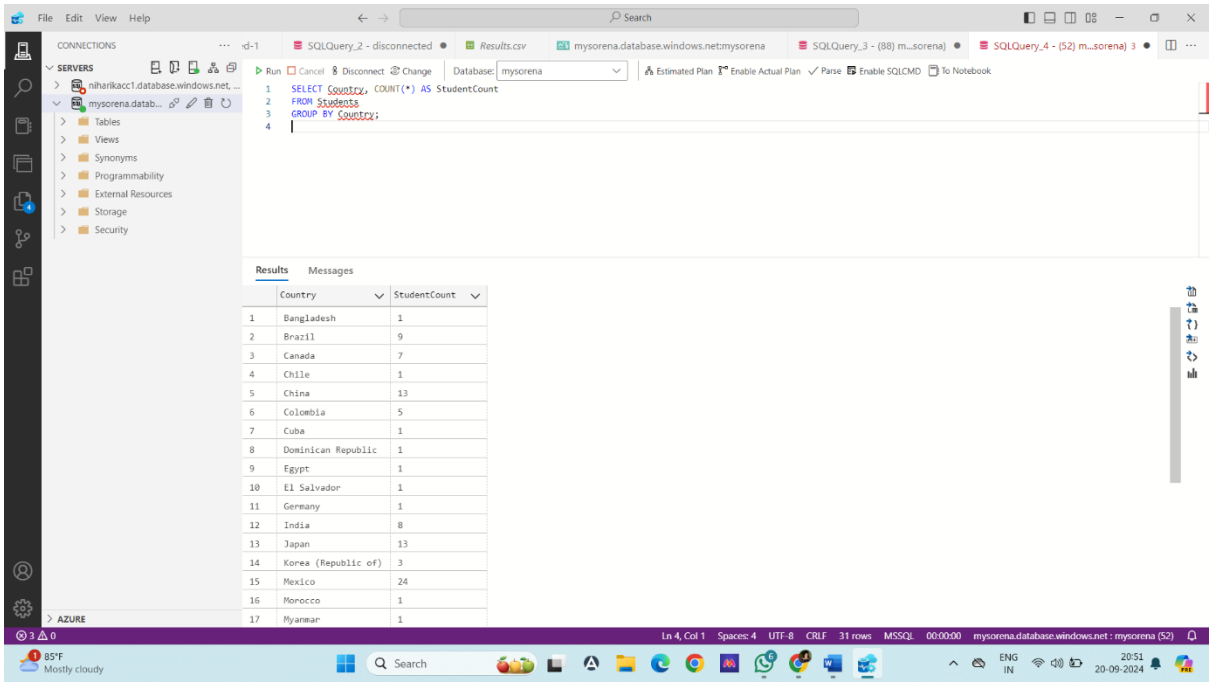

Microsoft Azure
Data Factory
mysorena
Search factory and documentation
mysorena@mail.uc.edu
UNIVERSITY OF CINCINNATI
Thinking of migrating to Fabric? Already made the switch? We want to hear from you! Spend just 20 minutes on our Fabric migration survey to shape the future of Data Factory, Synapse, and Spark in Fabric.
Data Factory
Validate all
Publish all
Successfully created
Successfully created BlobStorageLinkedService (Linked service).
General
Factory settings
Connections
Linked services
Integration runtimes
Microsoft Purview
Source control
Git configuration
ARM template
Author
Triggers
Global parameters
Data flow libraries
Security
Credentials
Customer managed key
Outbound rules
Linked services
Linked service defines the connection information to a data store or compute. Learn more
New
Filter by name
Annotations : Any
Showing 1 - 1 of 1 items
Name
Type
Related
Annotations
BlobStorageLinkedService
Azure Blob Storage
0
Search
ENG
IN
20:29
20-09-2024

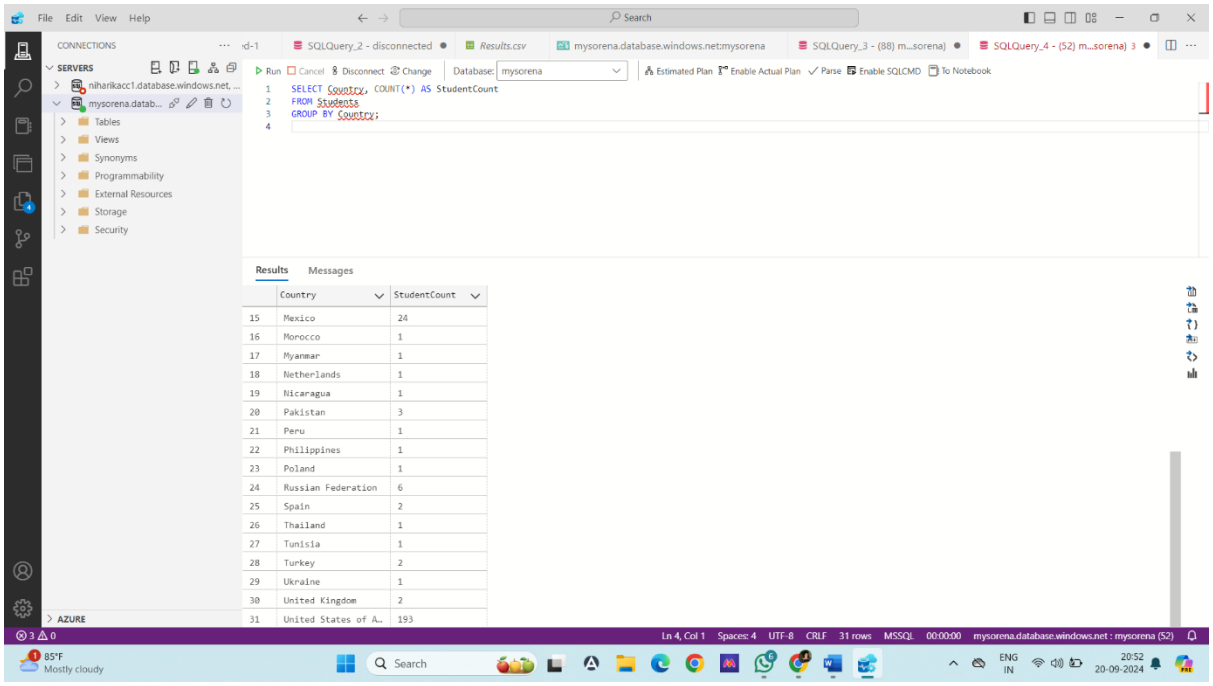

Microsoft Azure
Data Factory
mysorena
Search factory and documentation
mysorena@mail.uc.edu
UNIVERSITY OF CINCINNATI
Thinking of migrating to Fabric? Already made the switch? We want to hear from you! Spend just 20 minutes on our Fabric migration survey to shape the future of Data Factory, Synapse, and Spark in Fabric.
Data Factory
Validate all
Publish all
Successfully created
Successfully created AzureSQLLinkedService (Linked service).
Linked services
Linked service defines the connection information to a data store or compute. Learn more
New
Filter by name
Annotations : Any
Showing 1 - 2 of 2 items
Name
Type
Related
Annotations
AzureSQLLinkedService
Azure SQL Database
0
BlobStorageLinkedService
Azure Blob Storage
0
General
Factory settings
Connections
Linked services
Integration runtimes
Microsoft Purview
Source control
Git configuration
ARM template
Author
Triggers
Global parameters
Data flow libraries
Security
Credentials
Customer managed key
Outbound rules


Microsoft Azure
Data Factory
mysorena
Search factory and documentation
mysorena@mail.uc.edu
UNIVERSITY OF CINCINNATI
Thinking of migrating to Fabric? Already made the switch? We want to hear from you! Spend just 20 minutes on our Fabric migration survey to shape the future of Data Factory, Synapse, and Spark in Fabric.
Data Factory
Validate all
Publish all
Preview experience
Off
Factory Resources
Filter resources by name
Pipelines
mysorena
Change Data Capture (preview)
Datasets
SQLStudentTable
StudentCSV
Data flows
Power Query
StudentCSV
SQLStudentTable
mysorena
Activities
Move and transform
Copy data
Azure Data Explorer
Azure Function
Batch Service
Custom
Databricks
Notebook
Jar
Python
Data Lake Analytics
Validate
Validate copy runtime
Debug
Add trigger
Copy data1
Properties
General
Related
Name *
mysorena
Description
Annotations
New
General
Source
Sink
Mapping
Settings
User properties
Name *
Copy data1
Learn more
Description
Activity state
Activated
Deactivated
Timeout
0.12:00:00


Microsoft Azure
Data Factory
mysorena
Search factory and documentation
mysorena@mail.uc.edu
UNIVERSITY OF CINCINNATI
Thinking of migrating to Fabric? Already made the switch? We want to hear from you! Spend just 20 minutes on our Fabric migration survey to shape the future of Data Factory, Synapse, and Spark in Fabric.
Validate all
Publish all 3
Preview experience
Off
Factory Resources
Filter resources by name
Pipelines 1
mysorena
Change Data Capture (preview) 0
Datasets 2
SQLStudentTable
StudentCSV
Data flows 0
Power Query 0
StudentCSV
SQLStudentTable
mysorena
Activities
Move and transform
Copy data
Azure Data Explorer
Azure Data Explorer C...
Azure Function
Azure Function
Batch Service
Custom
Databricks
Notebook
Jar
Python
Data Lake Analytics
Validate
Cancel options
Add trigger
Copy data
Copy data1
Properties
General
Related
Name *
mysorena
Description
Annotations
New
Parameters
Variables
Settings
Output
Pipeline run ID: 6483f7a6-61ee-4a8d-aebb-f6e60a53f2fb
Pipeline status In progress
All status
Monitor in Azure Metrics
Export to CSV
Showing 0 - 0 of 0 items
Activity name
Activity status
Activity type
Run start
High pollen
Tomorrow
Search
ENG
IN
20:42
20-09-2024


Microsoft Azure
Data Factory
mysorena
Search factory and documentation
mysorena@mail.uc.edu
UNIVERSITY OF CINCINNATI
Thinking of migrating to Fabric? Already made the switch? We want to hear from you! Spend just 20 minutes on our Fabric migration survey to shape the future of Data Factory, Synapse, and Spark in Fabric.
Validate all
Publish all
Publishing completed
Successfully published
Factory Resources
Filter resources by name
Pipelines 1
mysorena
Change Data Capture (preview) 0
Datasets 2
SQLStudentTable
StudentCSV
Data flows 0
Power Query 0
StudentCSV
SQLStudentTable
mysorena
Activities
Move and transform
Copy data
Azure Data Explorer
Azure Data Explorer C...
Azure Function
Azure Function
Batch Service
Custom
Databricks
Notebook
Jar
Python
Data Lake Analytics
Validate
Debug
Add trigger
Copy data
Copy data1
Properties
General
Related
Name *
mysorena
Description
Annotations
New
Parameters
Variables
Settings
Output
Pipeline run ID: 6483f7a6-61ee-4a8d-aebb-f6e60a53f2fb
Pipeline status Succeeded
View debug run consumption
All status
Monitor in Azure Metrics
Export to CSV
Showing 1 - 1 of 1 items
Activity name
Activity status
Activity type
Run start
Copy data1
Succeeded
Copy data
9/20/2024, 8:42:
Upcoming
Earnings
Search
ENG
IN
20:43
20-09-2024


SELECT Country, COUNT(*) AS StudentCount
FROM Students
GROUP BY Country;
Results
Messages
Country
StudentCount
1 Bangladesh 1
2 Brazil 9
3 Canada 7
4 Chile 1
5 China 13
6 Colombia 5
7 Cuba 1
8 Dominican Republic 1
9 Egypt 1
10 El Salvador 1
11 Germany 1
12 India 8
13 Japan 13
14 Korea (Republic of) 3
15 Mexico 24
16 Morocco 1
17 Myanmar 1


SELECT Country, COUNT(*) AS StudentCount
FROM Students
GROUP BY Country;
Results
Messages
Country
StudentCount
15 Mexico 24
16 Morocco 1
17 Myanmar 1
18 Netherlands 1
19 Nicaragua 1
20 Pakistan 3
21 Peru 1
22 Philippines 1
23 Poland 1
24 Russian Federation 6
25 Spain 2
26 Thailand 1
27 Tunisia 1
28 Turkey 2
29 Ukraine 1
30 United Kingdom 2
31 United States of A... 193

- **Authentication**: SQL Server Authentication
- **Server name**: (mysorena)-server.database.windows.net
- Database name: (mysorena)-db
- **Login**: SQL Server username (mysorena)
- **Password**: nmg_Blueberry
- Ensure the firewall settings and credentials are correct

Total Estimated Monthly Cost for Azure Student Subscription:

- Azure Storage Account: $0.018
- Azure SQL Database (Basic Tier): $4.99
- Azure Data Factory (Pipeline Activity Runs): $4
- Estimated Total Cost: Around $9 per month.

Project Overview

Objective: The main goal of this project was to create a complete data pipeline using Azure cloud services. We needed to pull data from a CSV file with student information, process it, and load it into an Azure SQL Database for analysis.

Tools Used:

- **Azure Subscription:** The foundation for hosting our cloud services.
- **Azure Storage Account:** Used to upload and store the student dataset CSV file.
- **Azure SQL Database:** A relational database for storing and querying the processed student data.
- **Azure Data Factory:** The primary tool for building the data pipeline to move data from Azure Storage into the SQL Database.
- **Azure Data Studio:** Used to query the database after uploading the data.

Implementation Steps:

1. **Setting Up Azure Storage Account:**
    - Created a storage account and a Blob Storage container.
    - Uploaded the student-dataset.csv file into Blob Storage.

2. **Setting Up Azure SQL Database:**
    - Set up the Azure SQL Database and adjusted the firewall settings to allow remote access.
    - Created a table called "students" to store the student information.

3. **Building the Data Pipeline with Azure Data Factory:**

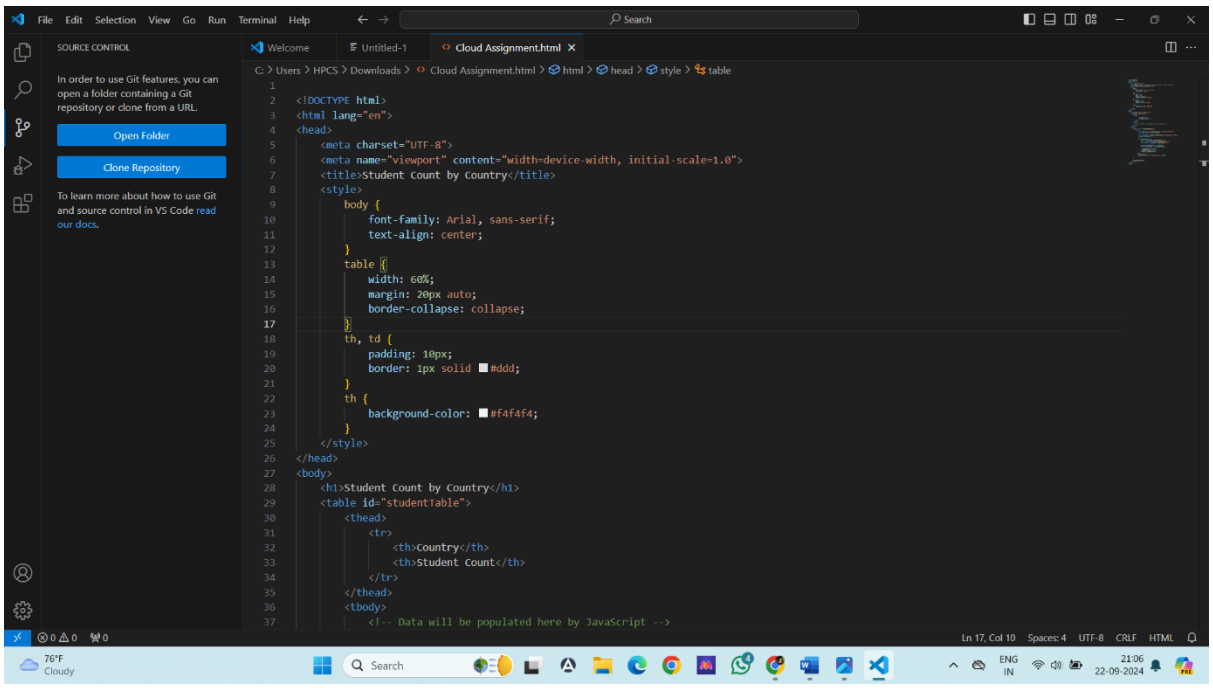
    - Created an instance of Azure Data Factory.
    - Set up linked services for both Azure Blob Storage and the Azure SQL Database.
    - Created two datasets: one for the CSV file and another for the SQL table where the data would go.
    - Built a pipeline that transfers data from the CSV file to the SQL Database and tested it to ensure everything worked smoothly.

4. **Testing and Querying:**

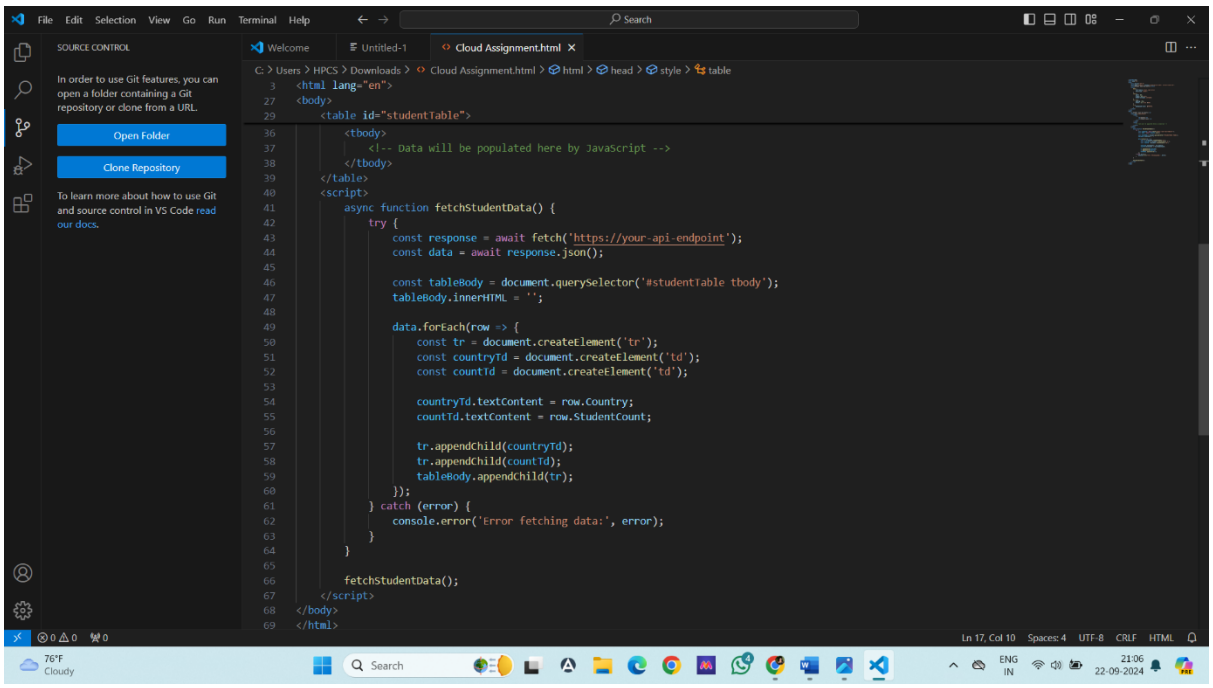
    - Connected to the Azure SQL Database using Azure Data Studio and checked that the data was correctly inserted into the "students" table.

EXTRA CREDIT

Develop a Web Application:



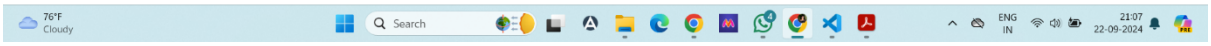
```
<!DOCTYPE html>
<html lang="en">
<head>
    <meta charset="UTF-8">
    <meta name="viewport" content="width=device-width, initial-scale=1.0">
    <title>Student Count by Country</title>
    <style>
        body {
            font-family: Arial, sans-serif;
            text-align: center;
        }
        table {
            width: 60%;
            margin: 20px auto;
            border-collapse: collapse;
        }
        th, td {
            padding: 10px;
            border: 1px solid #ddd;
        }
        th {
            background-color: #f4f4f4;
        }
    </style>
</head>
<body>
    <h1>Student Count by Country</h1>
    <table id="studentTable">
        <thead>
            <tr>
                <th>Country</th>
                <th>Student Count</th>
            </tr>
        </thead>
        <tbody>
            <!-- Data will be populated here by JavaScript -->
```

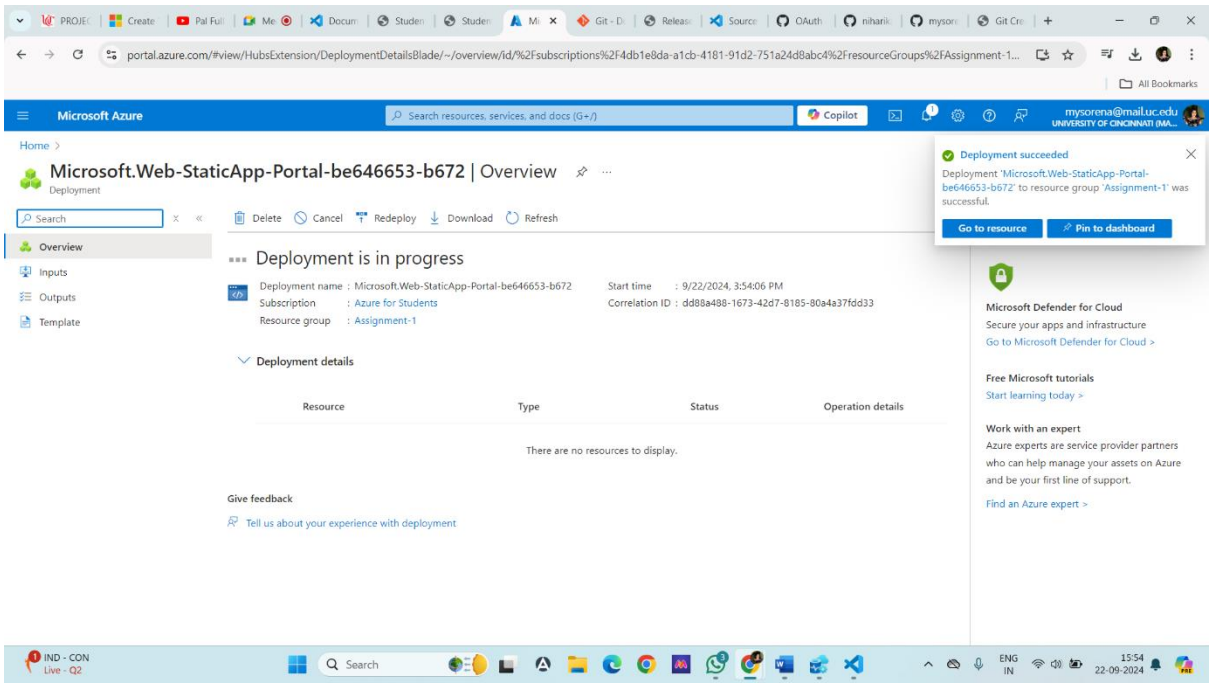


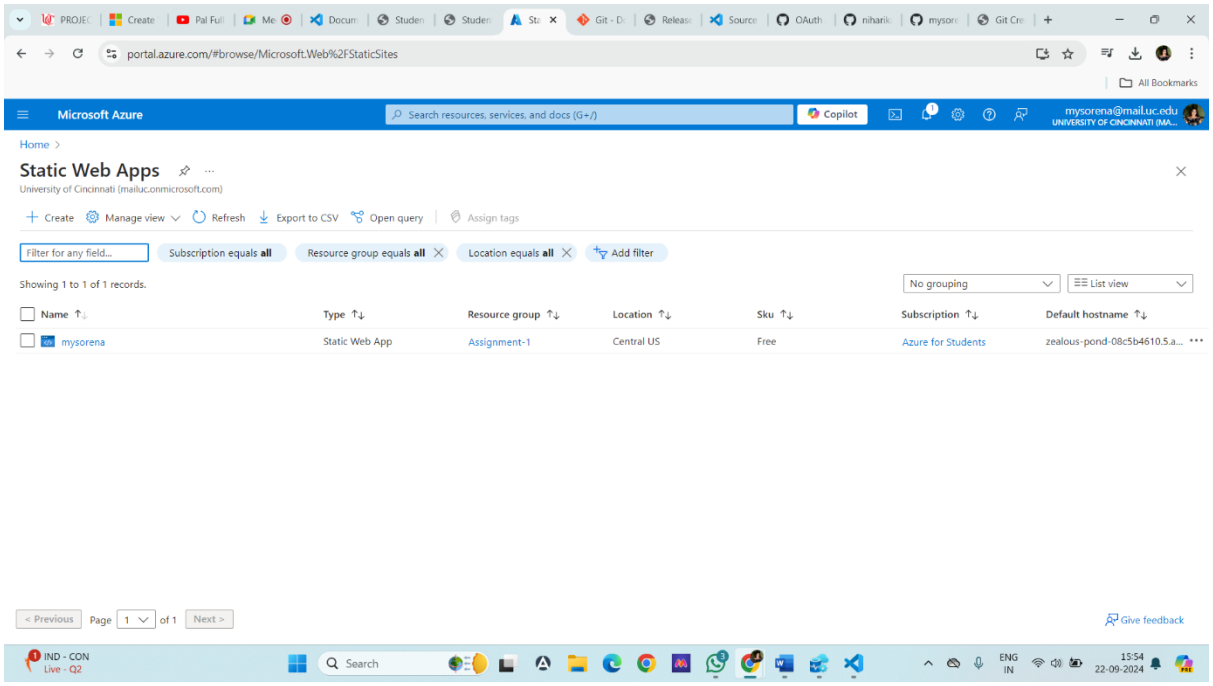
```
    <table id="studentTable">
        <tbody>
            <!-- Data will be populated here by JavaScript -->
        </tbody>
    </table>
    <script>
        async function fetchStudentData() {
            try {
                const response = await fetch('https://your-api-endpoint');
                const data = await response.json();

                const tableBody = document.querySelector('#studentTable tbody');
                tableBody.innerHTML = '';

                data.forEach(row => {
                    const tr = document.createElement('tr');
                    const countryTd = document.createElement('td');
                    const countTd = document.createElement('td');

                    countryTd.textContent = row.Country;
                    countTd.textContent = row.StudentCount;

                    tr.appendChild(countryTd);
                    tr.appendChild(countTd);
                    tableBody.appendChild(tr);
                });
            } catch (error) {
                console.error('Error fetching data:', error);
            }
        }

        fetchStudentData();
    </script>
</body>
</html>
```

Student Count by Country

Country	Student Count



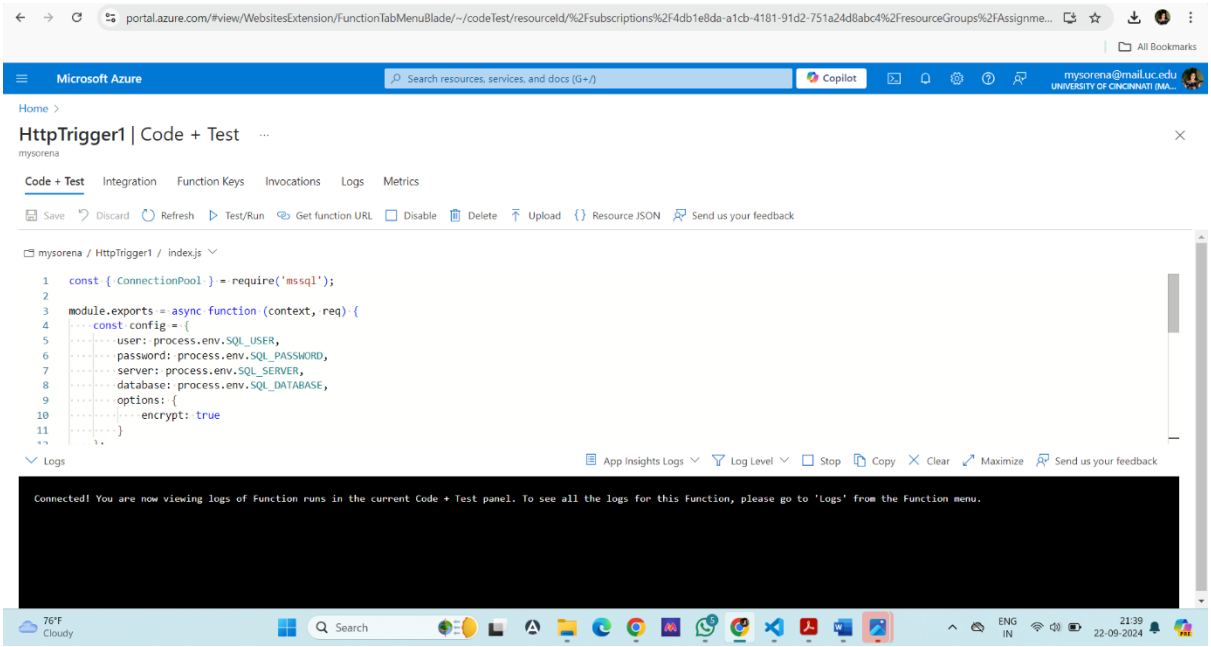


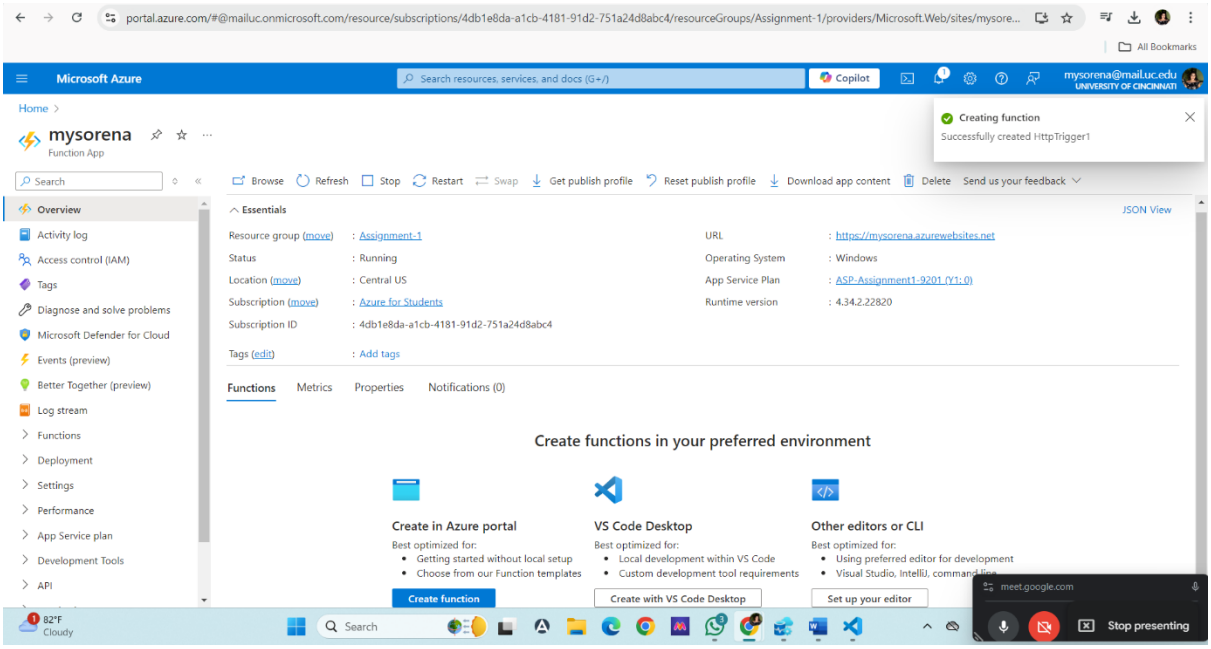


Repository Link:

https://github.com/niharikamg/mysorena

Create an API to Fetch Data from Azure SQL Database:





Student Count by Country

Country	Student Count
Bangladesh	1
Brazil	9
Canada	7
Chile	1
China	13
Colombia	5
Cuba	1
Dominican Republic	1
Egypt	1
El Salvador	1
Germany	1
India	8
Japan	13
Korea (Republic of)	3
Mexico	24
Morocco	1
Myanmar	1
Netherlands	1
Nicaragua	1
Pakistan	3
Peru	1

WEBSITE Link: https://mysorena.azurewebsites.net

Student Count by Country

Country	Student Count
Bangladesh	1
Brazil	9
Canada	7
Chile	1
China	13
Colombia	5
Cuba	1
Dominican Republic	1
Egypt	1
El Salvador	1
Germany	1
India	8
Japan	13
Korea (Republic of)	3
Mexico	24
Morocco	1
Myanmar	1
Netherlands	1
Nicaragua	1
Pakistan	3
Peru	1

Objective:

The goal of this project is to develop a static web application, hosted on Azure, that retrieves and displays student count data by country. The student data is stored in an Azure SQL Database, and the web app connects to an API that queries the database to present the results.

Tools Used:

1. Azure SQL Database: Stores the student data, including student IDs, names, and countries.
2. Azure Data Factory: Builds a pipeline for processing and transferring data to the SQL Database.
3. Azure Storage: Hosts the CSV file with student data, which is processed and transferred to the database.
4. Azure Static Web App: Hosts the front-end application that shows the student count by country.
5. Azure Functions: Provides the API backend to retrieve the student count data from the SQL Database.